Randomization

Michael Proschan, Ph.D.
Mathematical Statistician
National Heart, Lung, and Blood Institute

Outline

- Introduction
- Common types of randomization
  - Simple (complete) randomization
  - Random permuted blocks
  - Stratified randomization
- Other methods
  - Biased coin design
  - Adaptive randomization & minimization
  - Response-adaptive methods
- A true horror story

- Randomization to treatments separates clinical trials from all other studies; don’t muck it up!

- Randomization
  - eliminates selection bias
  - balances arms with respect to prognostic variables (known and unknown)
  - forms basis for statistical tests

- E.g., suppose 1000 women;

Expected & worse case allocation across T and C:

	% assigned to control	% assigned to treatment
Expected	50%	50%
95% extremes:	47%	53% or
	53%	47%

- Randomization considered so important that the ***Intention-to-treat (ITT) principle*** considered sacrosanct: ***Analyze by treatment randomized to irrespective of compliance***
  - If patient assigned to bypass surgery refuses surgery, still counted in bypass arm
  - That way compare comparable groups

- Otherwise groups may not be comparable
- E.g., in trial comparing medicine to biofeedback:
  - no theoretical reason to think patients *complying* with biofeedback are comparable to patients *complying* with medicine
  - there is theoretical reason to think patients *randomized* to biofeedback are comparable to patients *randomized* to medicine
- Avoid missing data!

- Predecessor to randomization: Alternating assignments (TCTCTCTC…)
- Arrowsmith (Sinclair Lewis, 1925), page 387:

  “These unfortunate cases he treated, giving the phage to alternate patients,…”

- Problems with alternating:
  - No assurance of comparability
  - Unblinding one unblinds all

- First trial to randomize was tuberculosis trial Amberson (1931)
  - 12 pairs of patients
  - within each pair, flipped coin to see who received treatment
- Diehl (1938) thought to be first to randomize in parallel-arm trial, but in speech to University of Minnesota chapter of Sigma Chi:

“At the beginning of the study, students who volunteered to take these treatments were assigned *alternately* and without selection to control groups and experimental groups…”

Not randomization!

- How do you randomize?
- Could flip coin for each participant—called *complete randomization* or *simple randomization*
- Problem: can get imbalance in # Ts and Cs, especially in smaller trials
  – Imbalance in prognostic factors more likely
  – Inefficient for estimating treatment effect

- E.g., in trial of 10 participants, treatment effect variance for 5-5 split relative to 7-3 split is

$$(1/5+1/5)/(1/7+1/3)=.84$$

- 7-3 split only 84% as efficient as 5-5 split

(T,C) Imbalance with 10 Participants

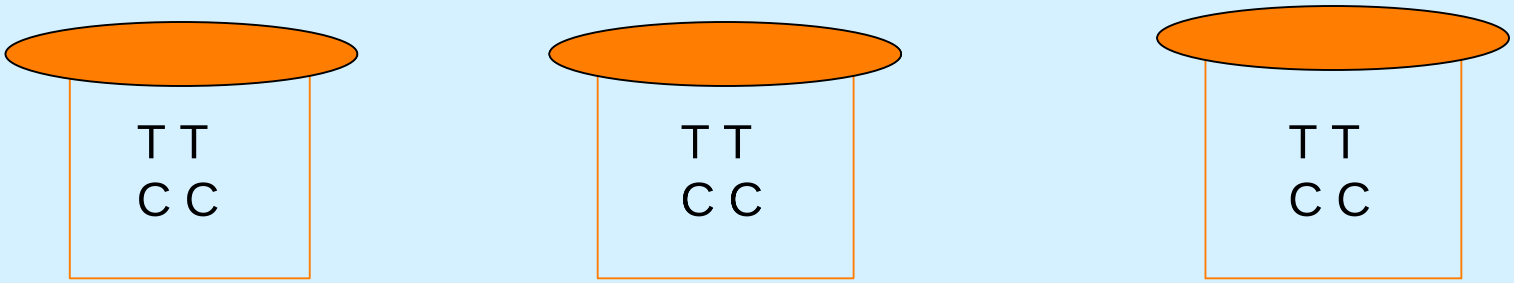
(#T, #C)	Probability	Efficiency
(5,5)	.246	1
(4,6) or (6,4)	.410	.96
(3,7) or (7,3)	.234	.84
(2,8) or (8,2)	.088	.64
(1,9) or (9,1)	.020	.36
(0,10) or (10,0)	.002	0

- Even if treatment balanced at end of trial, may be unbalanced at some time
- E.g., may be balanced at end with 400 participants, but first 10 might be

  CCCCTCTCTC
- Because we monitor trials over time, we want balance over time

Random Permuted Blocks

- To balance over time, could randomize in blocks (called *random permuted blocks*)
- Conceptually, for blocks of size 4: put 2 T labels & 2 C labels in hat: for next 4 participants, draw labels at random without replacement from hat
- TTCC TCTC TCCT CTTC CTCT CCTT all equally likely



T CT C   C CT T   C T C C

Pts 1 2 3 4   5 6 7 8   9 10 11 12…

Forces balance after every 4

- The smaller the block size, the more often balance is forced: e.g., in trial of 100,
  - blocks of size 2 force balance after every 2
  - A block of size 100 forces balance only at end
- From accidental bias/efficiency standpoint, more balance is good
- From selection bias standpoint, more balance is bad (in unblinded trial)

- E.g., with blocks of size 2 in unblinded trial, I know every second participant’s assignment **in advance**
- I can veto potential participants until I find one I like (sick one if next assignment is control, healthy one if next patient is treatment)
- Bigger issue in behavioral trials because of difficulty in blinding

- Even with larger blocks, in unblinded trial you know some assignments in advance
- E.g., with blocks of size 8 if first 5 are TCTTCT, know next 2 are C
- Using more than 1 block size makes it harder to guess
- But don’t make one block size a multiple of the other because then know where blocks could start

- E.g., with block sizes 4 and 8, blocks can only start at a multiple of 4
- If see TCCTCCTCC

  balanced after 4 but not 8, so first block had to be size 4, second had to be size 8;

(TCCT) (CCTCC _ _ _)

Know next 3 assignments are T

- Make it harder to guess next assignment
  - Don’t tell investigators block size
  - Use more than 1 block size (e.g., 6 and 8)
  - Do not make one block size a multiple of the other
  - In fact, make the greatest common divisor of the block sizes 2

- Sometimes want to balance treatment assignments within subgroups
- Especially important if subgroup size is small
- E.g., with 6 diabetics, if use complete randomization, there is 22% chance of 5-1 or 6-0 split!

- To avoid this problem could *stratify* the randomization (use blocked randomization separately for diabetics & nondiabetics)
- E.g., for blocks of size 6,

Diabetics	Nondiabetics
CTTCCT	TTCTCC TCCTTC…

Other Randomization Schemes

- Permuted block & stratified randomization most popular methods in clinical trials, but sometimes other methods used
- With Efron’s biased coin design, flip fair coin until there is a treatment imbalance, then flip unfair coin with probability 2/3 for under-represented treatment

- Efron's biased coin design

Step	0	1	2	3	4…
P(T)	1/2	2/3	1/2	1/3	1/2…
Actual assignment	C	T	T	C	T…

- Competitor of permuted block randomization
- Advantage: Can never be *sure* of next assignment

- Other methods compete with stratified randomization to balance prognostic factors—*adaptive randomization* and *minimization*
- Idea: Measure total imbalance through an imbalance function; rig it so next assignment more likely to reduce imbalance

- E.g., suppose have factors gender and race, & so far:

	Gender (G)		Hypertension (H)	
	M	F	Yes	No
T	10	3	8	5
C	8	3	6	5

I=2x(G imbalance)+3x(H imbalance)

Next patient is male, non hypertensive

	Gender (G)		Hypertension (H)	
	M	F	Yes	No
T	10	3	8	5
C	8	3	6	5

I=2x(G imbalance)+3x(H imbalance)

If next patient is T, I=2x(11-8)+3x(6-5)=9

If next patient is C, I=2x(10-9)+3x(6-5)=5

Flip unfair coin with P(C)=2/3

- *Minimization* uses same idea but eliminates almost all randomization
- Assign next patient to minimize imbalance function
- Only use randomization if get same imbalance whether next patient assigned to T or C

- Advantage of adaptive randomization over stratified randomization: Can’t stratify on many factors
- E.g., in extreme, so many strata that each contains only 1 participant
- Then stratified randomization equivalent to flipping coin for each participant—same as complete randomization

- Disadvantage of adaptive randomization and biased coin design: Analyzing data not as straightforward as for permuted block design

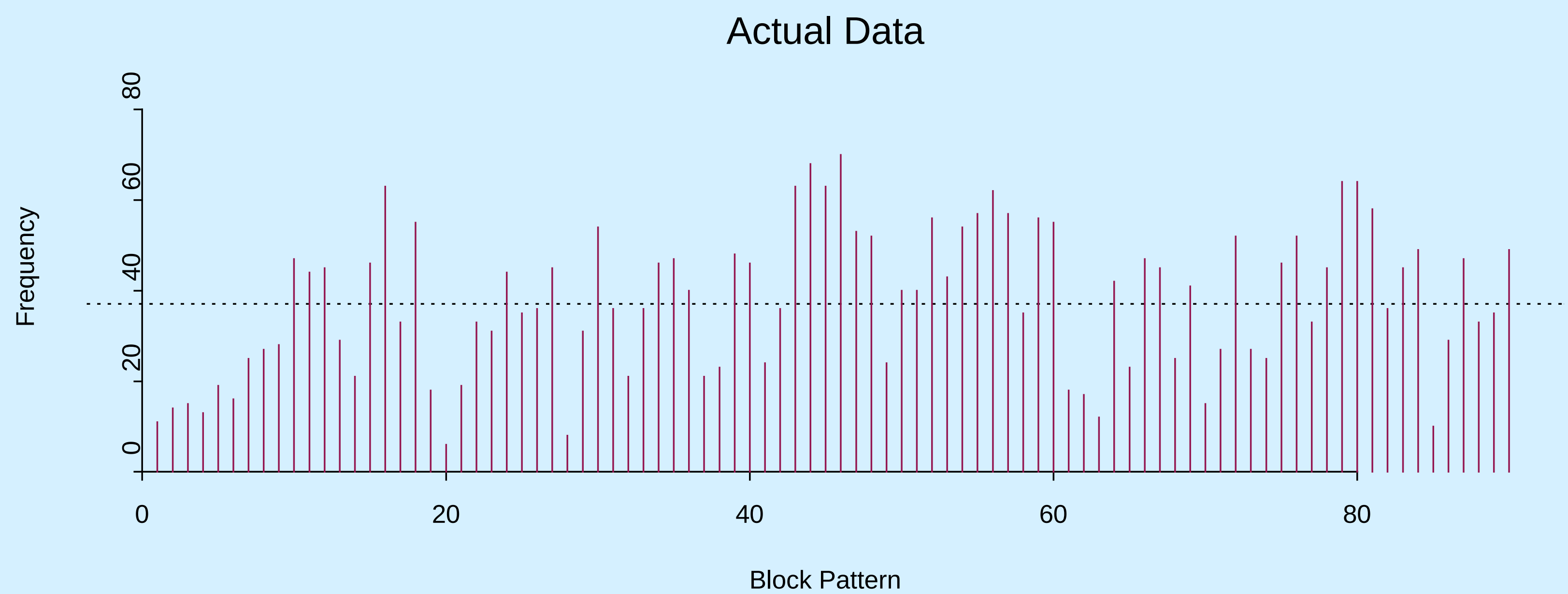
- Analyze as you randomize principle

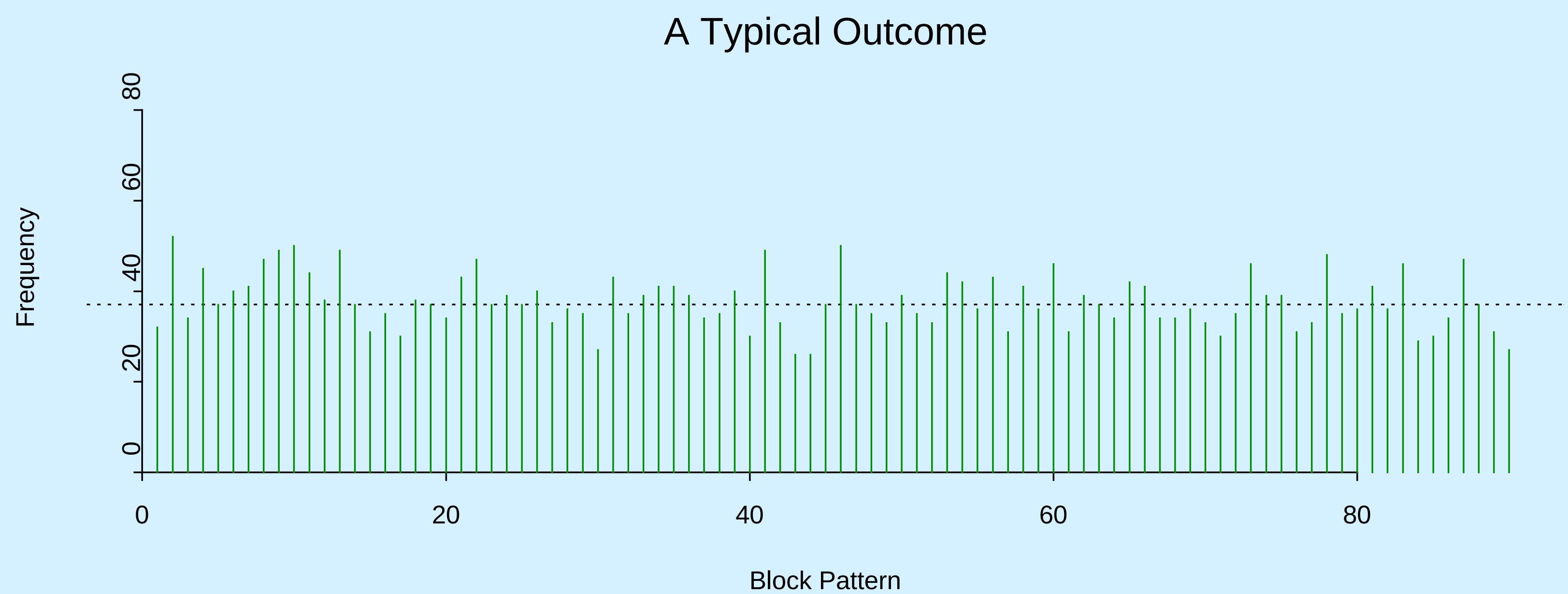
- Other methods even scarier
- E.g., response-adaptive designs change probabilities based on results of previous patients
- Even more of a nightmare to analyze (ECMO)

A Real Horror Study

- Three-armed trial
- Arms 2 and 3 were different doses of same drug
  - started at same dose; later, arm 3 to ramp up
- At interim look at data, many more adverse events in arm 2
- Problem: At time of interim look, arms 2 and 3 were at exactly same dose!

- The statistician investigated the randomization
- Baseline characteristics similar across arms
- Looked at blocked randomization
  - Used blocks of size 6, e.g., 112233, 121233,…332211
  - 90 different block patterns, expect about 37 participants per pattern


Actual Data
Frequency
0
20
40
60
80
0
20
40
60
80
Block Pattern


A Typical Outcome
Frequency
0
20
40
60
80
0
20
40
60
80
Block Pattern

- A disproportionately high percentage of participants assigned to blocks beginning with 22
- Protocol was complicated & there was a learning curve; many adverse events occurred early in trial
- Early randomizations were more likely to be treatment 2
- Trial was ruined!

Summary

- Randomization separates clinical trials from other studies
  - Tends to balance arms with respect to prognostic factors
  - Eliminates selection bias
  - guarantees validity of statistical tests
- Don’t jeopardize the randomization!
  - Follow the intention-to-treat principle
  - Avoid missing data

Summary (continued)

- To achieve balance of Ts and Cs:
  - Random permuted blocks most popular
    - The greater the balance, the better in terms of accidental imbalance, but worse in terms of selection bias
    - Use more than one block size
  - Biased coin design achieves balance but makes it impossible to be sure of next assignment

- To achieve covariate balance
  - Stratified randomization most popular
  - Adaptive randomization (and minimization) achieve better balance, but might pose technical difficulties in analysis

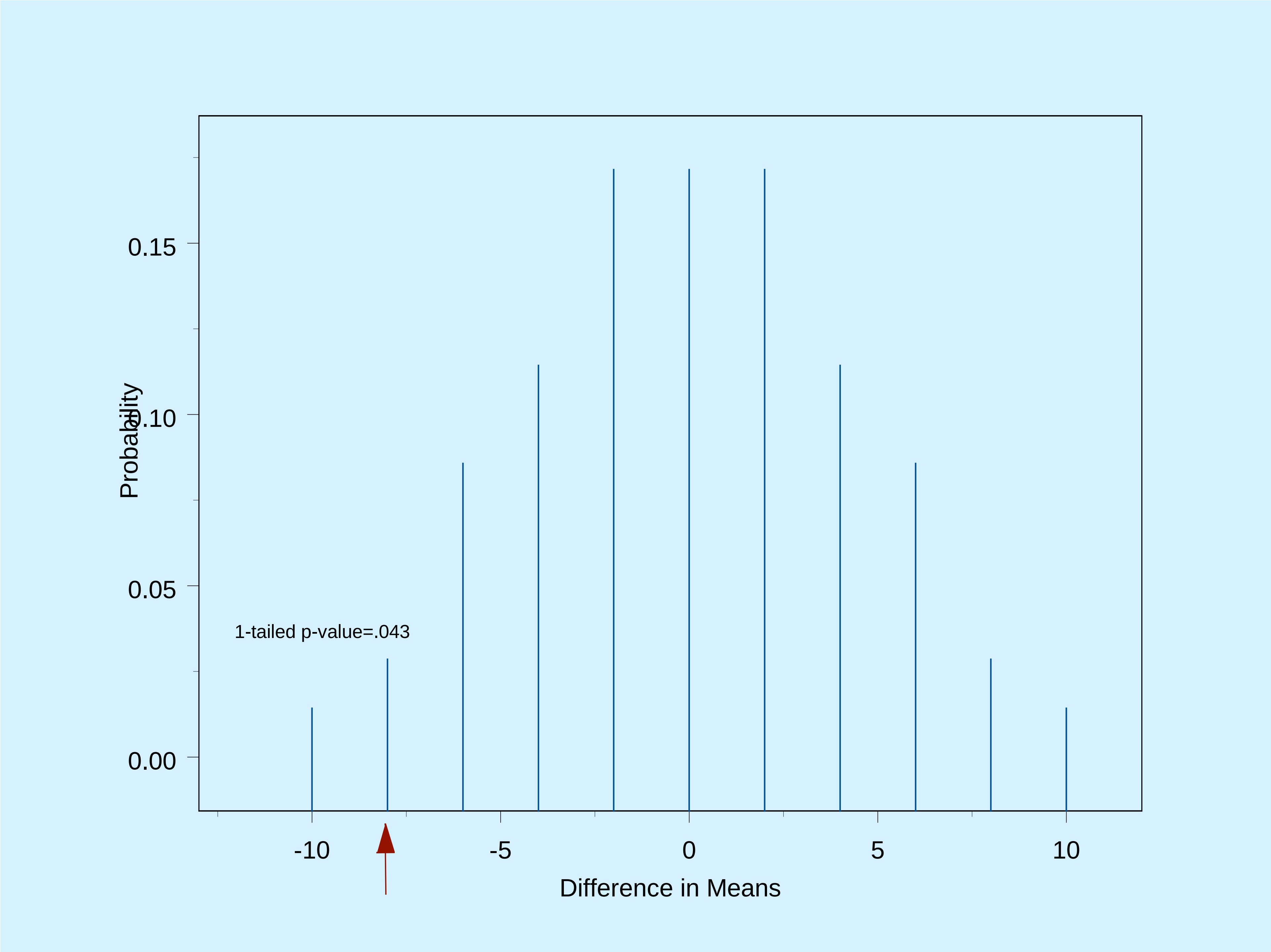
Appendix: A Permutation Test

-8	-8	-4	-4	0	4	4	8
T	T	T	C	T	C	C	C

$$(\text{Mean})_T - (\text{Mean})_C = -8$$

Scramble T,C labels, recompute difference in means, & repeat 1000s of times

Then see how far out -8 is in the tail of the "permutation distribution"


0.15
0.10
0.05
0.00
Probability
1-tailed p-value=.043
-10
-5
0
5
10
Difference in Means